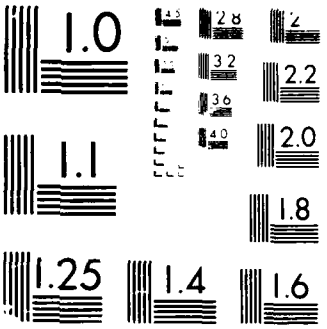
AD-A190 321 INFORMATION AND CODING CAPACITIES OF MISMATCHED GAUSSIAN CHANNELS(U) NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF STATISTICS C R BAKER SEP 87 LISS-19 1/1

UNCLASSIFIED N00014-86-K-0039 F/G 12/9 NL



MICROCOPY RESOLUTION TEST CHART

(4)
FILE COPY

AD-A190 321

Information and Coding Capacities of Mismatched Gaussian Channels

Charles R. Baker

Department of Statistics
University of North Carolina
Chapel Hill, NC 27514

DTIC
ELECTE
JAN 15 1988
S D
D

LISS 19
September, 1987

To appear: Proceedings First International Conference on Communication and Control Systems, Washington, D.C., June 18-20, 1987.

Research supported by ONR Contract N00014-86-K-0039

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

88 1 7 019

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA190321

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release: Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Department of Statistics; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): University of North Carolina, Chapel Hill, North Carolina 27514	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-86-K-0039
8c. ADDRESS (City, State and ZIP Code): Statistics & Probability Program, Arlington, VA 22217	10. SOURCE OF FUNDING NOS. — PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification): Information and Coding Capacities of . . .	
12. PERSONAL AUTHOR(S): C.R. Baker	
13a. TYPE OF REPORT: TECHNICAL; 13b. TIME COVERED: FROM _____ TO _____	14. DATE OF REPORT (Yr., Mo., Day): Sept., 1987; 15. PAGE COUNT: 9
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES — FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Gaussian channels; channel capacity; Shannon theory; Information Theory.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Recent results on coding capacity and information capacity for the mismatched Gaussian channel are discussed. Sufficient conditions for causal feedback to increase information capacity are given for the finite-dimensional discrete-time Gaussian channel.

TITLE CONT.: Mismatched Gaussian Channels

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED XX SAME AS RPT. XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: C.R. Baker	22b. TELEPHONE NUMBER (Include Area Code): (919) 962-2189; 22c. OFFICE SYMBOL

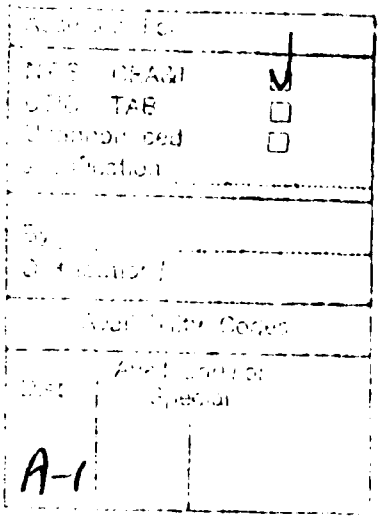
DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Abstract

Recent results on coding capacity and information capacity for the mismatched Gaussian channel are discussed. Sufficient conditions for causal feedback to increase information capacity are given for the finite-dimensional discrete-time Gaussian channel.



Introduction

The capacity (in the Shannon sense) of a communications channel is usually defined by either of two principal prescriptions. Information capacity is the supremum of the average mutual information between an input stochastic process (signal) and the noise-perturbed output process, with the supremum taken over an appropriate class of admissible input processes. The second definition is that of the supremum of all possible transmission rates, where the transmitted code words are subject to a constraint. For example, in the time-discrete additive channel, define the number of distinct code words transmitted by time t_n as $[e^{nR}]$, where $[x]$ is the integer part of x, and R is the "rate." If R is fixed and the maximum probability of decoding error goes to zero as $n \to \infty$ along some subsequence, then R is said to be an admissible rate (for the channel and the constraints). The (deterministic) coding capacity is then the supremum over all admissible rates. One can also consider random coding and other capacities connected with coding; only deterministic coding capacity will be considered here.

The additive Gaussian channel is a channel of primary practical importance. The received waveform is the sum of the transmitted waveform and a sample function from a Gaussian process: $Y = X + N$, where N is noise, X is signal. If the channel is without feedback, and X is a sample function from a stochastic process, then N is usually independent of X. With feedback, X will be a function of the past values of Y, and will thus depend upon N.

In this paper, a general discussion is first given of some recent results on information capacity and coding capacity of additive Gaussian channels when the constraint is mismatched to the channel noise; that is, the constraint is given in terms of a covariance that is different from that of the noise covariance. Such a "mismatched channel" is the usual case in practice, since one will rarely know the exact covariance of the noise. Moreover, in some situations, such as jamming channels, the mismatch occurs as an essential part of the problem formulation. The results given here on information capacity without feedback appear in [3]; the results on coding capacity will appear in [5]. A second set of new results summarized here consists of sufficient conditions for causal feedback to increase information capacity [2]. A statement and proof are given for the finite-dimensional time-discrete channel.

The channels considered here can be nonstationary and can have memory. Thus, in the discrete-time case, it is not required that the noise covariance matrix R_N be a diagonal matrix.

Information Capacity and Coding Capacity of Gaussian Channels Without Feedback

For information capacity of Gaussian channels without feedback, solutions are given in [1] and [3]. The framework there is for stochastic processes inducing measures on Hilbert space. These results can be extended to measures induced on a class of linear topological spaces; see [9] and [4].

Consider now the additive time-discrete Gaussian channel without feedback, with processes involved having sample paths in ℓ_2. Let X denote an input stochastic process, $Y = X + N$ as above, and $I[X, Y]$ the mutual information between X and Y (see, e.g., [1] for basic definitions). Let R_N denote the covariance operator of the noise, and let R_W denote another covariance operator. Define the constraint on X by $\mathcal{E}\|X\|_W^2 \leq P$, where $\mathcal{E}(\cdot)$ denotes expectation with respect to the probability on ℓ_2 defined by X, and $\|\cdot\|_W$ is the reproducing kernel Hilbert space (RKHS) norm for R_W: $\|y\|_W^2 = \|R_W^{-\frac{1}{2}}y\|^2$ ($\|\cdot\|$ the ℓ_2 norm); one can assume WLOG that R_W^{-1} exists. If $R_W = R_N$, then the supremum of $I(X, X+N)$ over all such admissible X processes is equal to $P/2$.

For the same channel, with deterministic coding used, for each $n \geq 1$, constrain each code word x to belong to $\mathbb{R}^n$ and to satisfy $\|x\|_{W,n}^2 \leq nP$, where $\|\cdot\|_{W,n}$ is the RKHS norm of R_W, with R_W^n the $n \times n$ matrix given by $R_W^n(ij) = R_W(ij)$, $i,j \leq n$. A code $\{k,n,\epsilon_n\}$ is then a set of k code words, each obeying the constraint, with maximum probability of decoding error being $\leq \epsilon_n$. A real number $R \geq 0$ is then an admissible rate if there exists a sequence $(\{[e^{nR}],n,\epsilon_n\})$ of codes such that $\epsilon_n \to 0$ as $n \to \infty$ along some subsequence. The supremum of all admissible rates is the coding capacity, denoted here by $C_W^\infty(P)$. If $R_W = R_N$, then $C_W^\infty(P) = \frac{1}{2}\log[1+P]$.

Those familiar with the Shannon theory will recognize the similarity of the above results to those obtained for the classical white noise channel with

a pure power constraint [6]. However, this similarity disappears when one examines the "mismatched" channel: $R_W \neq R_N$. The expression for the information capacity then takes one of several forms, depending on the relationship between R_W and R_N. For finite information capacity, one must have $R_N = R_W^{\frac{1}{2}}(I+S)R_W^{\frac{1}{2}}$, where I is the identity in ℓ_2 and S is a self-adjoint operator in ℓ_2 such that $(I+S)^{-1}$ exists and is bounded [3]. The information capacity then depends on the spectrum of S; specifically, on the smallest limit point of the spectrum, denoted by θ, and those eigenvalues (if such exist) of S that are strictly less than θ. See [3] or [4] for the various expressions. These expressions are considerably more complicated than that for the matched channel.

For coding capacity when $R_W \neq R_N$, one again obtains a rather complicated expression for the capacity. In [5], a solution is given for capacity under the assumption that S has a pure point spectrum. The solution is a function of the limit points of the spectrum of S and of their "relative importance." For the memoryless channel, where R_N is diagonal, this "relative importance" can be roughly described as the relative frequency of each limit point.

In the case where the spectrum of S has a single limit point, θ, and S has no eigenvalues strictly less than θ, one obtains a result analogous to that of the matched channel ($R_W = R_N$): the information capacity is equal to $\frac{1}{2}\frac{P}{1+\theta}$, and the coding capacity is equal to $\frac{1}{2}\log[1 + \frac{P}{1+\theta}]$.

In the analogous problems for the time-continuous channel, the constraint on the code words is given by $\|x\|^2_{W,T} \leq PT$, where x is required to belong to $L_2[0,T]$, and $\|\cdot\|_{W,T}$ is the RKHS norm of $R_{W,T}$. $R_{W,T}$ is obtained from a covariance function r_W, defined on $[0,\infty)\times[0,\infty)$, and $R_{W,T}$ is the integral operator defined by the restriction of r_W to $[0,T]\times[0,T]$. In this case, assuming that range($R_{W,T}$) is infinite-dimensional for some $T > 0$, the coding capacity when $R_{W,T} = R_{N,T}$ for all $T > 0$ is given by P/2. When $R_{W,T} \neq R_{N,T}$, with $R_{N,T} = R_{W,T}^{\frac{1}{2}}(I_T+S_T)R_{W,T}^{\frac{1}{2}}$, I_T the identity in $L_2[0,T]$, then the coding capacity depends on the behavior of $\{\theta_T,\ T > 0\}$ and $\{\lambda_n^T,\ n \geq 1,\ T > 0\}$, where θ_T is the

smallest limit point of the spectrum of S_T and $\{\lambda_n^T,\ n \geq 1\}$ is the set of eigenvalues of S_T that are strictly less than θ_T. If $\{\lambda_n^T,\ n \geq 1\}$ is empty for all sufficiently large T, then the coding capacity is $\frac{1}{2}\frac{P}{1+\theta}$, where $\theta = \lim_{T\to\infty} \theta_T$. However, in general $\frac{1}{2}\frac{P}{1+\theta}$ is only a lower bound for the coding capacity

Thus, the results for coding capacity and for information capacity of the mismatched channel ($R_W \neq R_N$) both differ significantly from the corresponding results for the matched channel. For further details, reference is made to [1], [3], and [5].

All of the above discussion is for the additive Gaussian channel without feedback. In the case of channels with causal feedback, the solutions for information capacity and for coding capacity have not been obtained in the case of the mismatched channel. For the matched channel, information capacity when N is the Wiener process has been obtained [8], and this has been extended to obtain capacity for some more general Gaussian processes [7]. In both cases, it has been found that causal feedback does not increase capacity. A solution has not been published for the general additive Gaussian channel, even for the matched case ($R_W = R_N$).

Feedback Capacity

Information capacity of the mismatched Gaussian channel with feedback is an open problem. It has long been speculated that causal feedback can increase capacity over the no-feedback situation. An answer will be given here to these questions for the discrete-time finite-dimensional channel; processes take values in $\mathbb{R}^K$. These results and other results for infinite-dimensional channels were announced at the 1986 IEEE Symposium on Information Theory [2].

The channel output is $Y = X - BY + N$, where X is the message process, N is Gaussian noise independent of the message, and B is a strictly-lower-triangular (SLT) matrix ($b_{ij} = 0$ for $j \geq i$). The transmitted signal is $X - BY$. All processes are defined on a probability space (Ω, β, μ), and $\mathcal{E}$ will be used to denote expectation with respect to μ. The capacity problem is the following:

$$\text{maximize } I[X, Y] \qquad \text{subject to } \mathcal{E}\|X - BY\|^2 \le P,$$

where $\|\cdot\|$ is the norm for a K-dimensional Euclidean space: $\|X\|^2 = \sum_{i=1}^{K} X_i^2$.

I[X, Y] denotes mutual information of X and Y. See [1] for definitions.

Let

$$C_F(P) = \sup_F I[X, X-BY+N]$$

$$C(P) = \sup_{F_1} I[X, X+N]$$

where $F = \{(X,B): \mathcal{E}\|X - BY\|^2 \le P, Y = X - BY + N, B \text{ SLT}\}$

$F_1 = \{X: \mathcal{E}\|X\|^2 \le P\}$.

An "elementary vector" in $\mathbb{R}^K$ is a vector x such that $x_k = 1$, $x_i = 0$ for $i \ne k$, some k in $\{1,2,\ldots,K\}$.

The main results of this section are contained in the following theorem.

THEOREM. $C_F(P) > C(P)$ for all $P > 0$ if the eigenmanifold for the smallest eigenvalue of R_N does not have a basis consisting entirely of elementary vectors which are eigenvectors of R_N.

$C_F(P) > C(P)$ for all sufficiently large P if R_N is not a diagonal matrix.

□

In order to prove the result, the problem will first be reformulated into an equivalent no-feedback problem involving a pure power constraint.

Reformulation of the Problem

$Y = X - BY + N$; since B is SLT, $Y = (I+B)^{-1}(X+N)$. Moreover, as $(I+B)^{-1}$ is 1:1, $I[X, Y] = I[X, X+N]$. The constraint is $\mathcal{E}\|X - BY\|^2 \le P$, which can be written as $\mathcal{E}\|X - B(I+B)^{-1}(X+N)\|^2 \le P$. Since B is SLT, I + B is lower triangular, so $(I+B)^{-1}$ is lower triangular and $B(I+B)^{-1}$ is again SLT. Given any SLT C, there exists a SLT B satisfying $C = B(I+B)^{-1}$; simply, $B = (I-C)^{-1}C$. The original feedback problem is thus equivalent to finding sup I[X, X+N] subject to $\mathcal{E}\|X - C(X+N)\|^2 \le P$, where C is any SLT matrix.

Using the above, attention can now be restricted to the following problems.

$$C_F(P) = \sup_{F'(P)} I[X, X+N]$$
$$C(P) = \sup_{F_1'(P)} I[X, X+N]$$

where $F'(P)$ is the set of all Gaussian random vectors in $\mathbb{R}^K$ such that $\mathcal{E}\|X - B(X+N)\|^2 \leq P$ for some SLT matrix B, and $F_1'(P)$ is the set of all Gaussian random vectors in $\mathbb{R}^K$ such that $\mathcal{E}\|X\|^2 \leq P$.

Structure of the Reformulated Problem

Let $H(\mathbb{R}^K,\mu)$ be the set of all K-component real random vectors f on (Ω,β) such that $\mathcal{E}\sum_{n=1}^{K} f_n^2(\omega) < \infty$. $H(\mathbb{R}^K,\mu)$ is a Hilbert space under the inner product $(f,g)_\mu = \mathcal{E}\sum_{n=1}^{K} f_n(\omega)g_n(\omega)$. Suppose that X and N are two mutually independent zero-mean Gaussian (w.r.t. μ) random vectors: $\mathcal{E}X_n(\omega)N_m(\omega) = 0$ for all $n,m \leq K$. Suppose also that N has non-singular covariance matrix R_N. Let $H_-(X+N)$ be the set of all random vectors f in $H(\mathbb{R}^K,\mu)$ having the form $f = B(X+N)$, where B is an SLT matrix. It is clear that $H_-(X+N)$ is a linear manifold. It is also closed in $H(\mathbb{R}^K,\mu)$ norm since

$$\|B^n(X+N) - B^m(X+N)\|_\mu^2 = \|(B^m-B^m)(X+N)\|_\mu^2$$
$$= \text{Trace } (B^n-B^m)(R_X+R_N)(B^n-B^m)^* \geq \gamma_0 \text{ Tr } (B^n-B^m)(B^n-B^m)^*,$$

where γ_0 is the minimum eigenvalue of R_N.

Thus, if $(B^n(X+N))$ is Cauchy in $H(\mathbb{R}^K,\mu)$, then Tr $(B^n-B^m)(B^n-B^m)^* \to 0$. This is equivalent to $\sum_{i,j=1}^{K} (B_{ij}^n-B_{ij}^m)^2 \to 0$. Hence (B_{ij}^n) must be Cauchy for each ij, and so the limit exists as an SLT matrix B.

Now let N be a fixed Gaussian vector. For any Gaussian vector X independent of N, let P_-X be the projection of X onto $H_-(X+N)$. The feedback problem is now to choose a Gaussian vector X so that I[X, X+N] is maximized, while $\|X - P_-X\|_\mu^2 \leq P$.

That is, if one chooses any Gaussian vector X with SLT feedback matrix B, such that $\mathcal{E}\|X - B(X+N)\|^2 \leq P$, then necessarily $\mathcal{E}\|X - B(X+N)\|^2 \geq \|X - P_-X\|_\mu^2$.

and since $P_X = C(X+N)$ for some SLT matrix C (since $H_(X+N)$ is closed) one can replace B with C and be assured that the constraint is still satisfied.

It can be seen from the above that $C_F(P) > C(P)$ if the optimum solution X for the no-feedback message is not orthogonal to $H_(X+N)$. In fact, if this condition is satisfied, then for the optimum no-feedback message X, and $\alpha \neq 0$, $\mathcal{E}\|\alpha X - B(\alpha X+N)\|^2 \leq P$ gives $\alpha^2 \mathcal{E}\|X\|^2 \leq P + \Delta$, where $\Delta = \text{Tr } B[\alpha^2 R_X + R_N]B^* = \alpha^2 \text{Tr } BR_X$ and $B(\alpha X + N)$ is the projection of αX onto $H_(\alpha X+N)$. Since $\mathcal{E}\|X\|^2 = P$ for the optimum no-feedback message X, setting $\alpha^2 \mathcal{E}\|X\|^2 = P + \Delta$ gives $\alpha^2 = 1 + \Delta/P$, so that $\alpha^2 > 1$ whenever $\Delta > 0$. Thus, one can replace X in the no-feedback problem with αX, use the upper bound $P + \Delta$ in place of P, and obtain a strict increase in capacity. Of course, Δ depends on α.

The above requires that the optimum no-feedback message X not be orthogonal to $H_(X+N)$. Since X is independent of N, this orthogonality condition occurs if and only if X is such that for <u>all</u> non-zero SLT matrices B, $\text{Tr } BR_X \neq 0$.

PROPOSITION. $\text{Tr } BR_X = 0$ for every SLT matrix B if and only if R_X is diagonal.

Proof. Since $(BR_X)_{ii} = \sum_{j<i} B_{ij} R_X(ji)$, it is clear that $\text{Tr } BR_X = 0$ for <u>every</u> SLT matrix B if R_X is diagonal. Now suppose that $\text{Tr } BR_X = 0$ for all SLT matrices B. For any $i,j \leq K$ such that $i > j$, choose the matrix B to be zero except for the ij component; then $\text{Trace } BR_X = b_{ij} R_X(ji) = 0$, so that $R_X(ji) = 0$. As R_X is symmetric, this shows that the condition $\text{Tr } BR_X = 0$ for all SLT matrices B implies R_X is diagonal. □

This development shows that feedback can increase capacity if the optimum no-feedback message X does not have uncorrelated components. From [3, Theorem 1], the optimum no-feedback signal covariance is given by

$$R_X = \frac{1}{J}\left[\sum_{i=1}^{J} \beta_i + P\right] \sum_{n=1}^{J} u_n u_n^* - \sum_{m=1}^{J} \beta_m u_m u_m^*,$$

where $\{u_n, n \leq K\}$ are o.n. eigenvectors of R_N corresponding to the increasing

sequence of eigenvalues $(1+\beta_n)$, and $J \leq K$ is the largest integer such that $P + \sum_{i=1}^{J} \beta_i \geq J\beta_J$. For all sufficiently small P, this gives $R_X = \frac{P}{L} \sum_{n=1}^{L} u_n u_n^*$, where L is the multiplicity of $1 + \beta_1$ as an eigenvalue of R_N. R_X will then not be diagonal if $\{u_i, i \leq L\}$ cannot be taken to consist of elementary vectors. If R_X is defined as above for $J > L$, then this property will again prevent R_X from being diagonal, since a diagonal R_X must have the K elementary vectors as a c.o.n. set of eigenvectors. This shows that $C_F(P) > C(P)$ for all $P > 0$ if the restriction of R_N to the eigenmanifold of $1 + \beta_1$ (as an eigenvalue of R_N) is not diagonal. Further, for all sufficiently large P,

$$R_X = \frac{1}{K} \left[\sum_{i=1}^{K} \beta_i + P + K \right] I - R_N.$$

This matrix is obviously non-diagonal if R_N is non-diagonal. These observations complete the proof of the Theorem. □

The above results give sufficient conditions for feedback to increase information capacity. It can be seen that the requirement that R_N not be diagonal is also a necessary condition if feedback is to increase capacity for some value of P, without assuming linear feedback.

Acknowledgement

The research summarized here was supported by ONR Contract N00014-86-K-0039.

References

1. C. R. Baker, Capacity of the Gaussian channel without feedback, *Inform. Contr.* 37 (1978), 70-89.

2. C. R. Baker, Information capacity of the Gaussian channel with feedback, Abs. of *Papers, IEEE Intern. Symp. on Inform. Theory*, Ann Arbor Mich., Oct. 6-9, 1986, 78.

3. C. R. Baker, Capacity of the mismatched Gaussian channel, to appear, *IEEE Trans. on Inform. Theory* (1987).

4. C. R. Baker, Information capacity of Gaussian channels, "Recent Advances in Communication and Control Theory," Eds. R.E. Kalman, G.I. Marchuk, A.E. Ruberti, and A.J. Viterbi, 2-15, Optimization Software, Inc., New York.

5. C. R. Baker, Coding capacity of the discrete-time Gaussian channel, to appear.

6. R. G. Gallager, "Information Theory and Reliable Communication," Wiley, New York (1968).

7. M. Hitsuda and S. Ihara, Gaussian channels and the optimal coding, *J. Multivariate Anal.*, vol. 5 (1975), 106-118.

8. T. T. Kadota, M. Zakai, and J. Ziv, Mutual information of the white Gaussian channel with and without feedback, *IEEE Trans. on Inform. Theory*, 17 (1971), 368-371.

9. I. W. McKeague, On the capacity of channels with Gaussian or non-Gaussian noise, *Inform. Contr.* 51 (1981), 153-173.

END
DATE
FILMED
5-88
DTIC